Regents Junior Faculty Fellowship Ryan Giordano

I propose to spend the summer of 2024 working on two collaborative research projects. The first, "neural network classifiers for Bayesian posteriors," promises to introduce a completely new set of Bayesian inference techniques based on ideas from simulation–based inference. The second, "black–box computable diagnostic weights for survey sampling," will bring a much–needed set of diagnostic tools to the vast majority of modern applied survey sampling. These two projects are different in scope — the first represents ground–breaking methodological research, and the second an application of my existing research to an urgent applied problem — but each rests on and contributes to my existing work and expertise on approximate Bayesian computation and sensitivity analysis.

Neural network classifiers for Bayesian posteriors

Bayesian statistical techniques are a conceptually powerful set of tools for representing and quantifying uncertainty, and are increasingly popular across the physical and social sciences. Often, a statistical analysis involves a single quantity of interest, such as the effect of a policy intervention [Meager, 2019], the type of an astronomical object [Regier et al., 2019], the outcome of an election [Gelman and Heidemanns, 2020], or the identity of an ancestral genetic population [Pritchard et al., 2000]. Bayesian statistics is able to propagate uncertainty from any unknown latent modeling quantities to the final estimate. But this conceptual strength is a computational weakness, since even approximately accounting for a large number of latent quantities is computationally intensive. Bayesian estimates often take hours to days to compute, and it is of considerable interest to develop computationally efficient, approximate Bayesian procedures [Blei et al., 2017, AABI, 2024].

In consultation with a staff scientist at LBNL, I have recently developed a new approach to Bayesian inference based on neural network classifiers (NNC). The idea is derived from a technique for point estimation in simulation–based inference (SBI),[1] a technique I will refer to as SBI-NNC [Cranmer et al., 2020]. Rather than learning a likelihood directly, SBI-NNC exploits the fact that optimal neural network classifiers learn likelihood ratios. I have shown that a variant of the SBI-NNC trick can be applied to learn Bayesian marginals without having to learn the distribution of all the latent variables, at the cost of training a NNC on a single classification task. I will refer to my technique as Bayes–NNC.

Both classical Bayesian procedures and the existing SBI-NNC trick are difficult to validate in practice, due to the lack of a computable ground truth. Amazingly, Bayes–NNC does not suffer from this shortcoming, and its accuracy is readily testable using simulation–based calibration (SBC) [Talts et al., 2018]. Put together, Bayes–NNC and SBC offer a way to learn Bayesian posterior densities of low–dimensional quantities of interest with strong, computable statistical accuracy guarantees. Interestingly, SBC is well–known but rarely used in practice, since it is typically computationally prohibitive to compute the posterior at many different datapoints. However, Bayes–NNC learns the posterior for many datasets simultaneously, permitting efficient use of SBC in practice.

To my knowledge, there are no existing Bayesian techniques that offer the advantages of Bayes-NNC and SBC. Bayesian approaches to simulation–based inference are not new, but

[1] That is, inference in problems without a tractable likelihood function.

existing techniques are built on high–dimensional density approximation, such as normalizing flows [Cranmer et al., 2020, Papamakarios et al., 2021]. As with other approximate inference techniques, this set of tools approximates the entire posterior, even when only a low–dimensional marginal is of interest. To the best of my (and my LBNL collaborator's) knowledge, Bayes–NNC is new, and offers a distinct and advantageous set of computational tradeoffs relative to existing Bayesian inference methods.

Black–box computable diagnostic weights for survey sampling

Most modern surveys — such as polling about the upcoming presidental election — must overcome the fact that their sampled population is different from the target population [Gelman, 2007]. For example, the set of people responding to an internet survey about political preferences is likely to differ systematically from the full population of voters, and it is extremely useful to be able to check that the re-weighting is accurate, for example by checking that key demographic variables are balanced by the re-weighting [Li et al., 2018, B. et al., 2021]. Unfortunately, the most accurate and most commonly used statistical procedures for inferring the polling responses of rare demographic groups are nonlinear, and so do not readily admit diagnostic weights [Gelman, 1997, 2007].

In collaboration with a UC Berkeley professor of public policy, I have shown that one can compute "local diagnoistic weights" for non–linear statistical procedure, provided a much–needed diagnostic that is currently unavailable. The local weights I derive are closely related to the classical "influence function" of robust statistics [Mises, 1947, Hampel et al., 1986, Giordano et al., 2019]. Though the influence function is well–studied in the frequentist literature, it has been relatively neglected in the Bayesian literature (with my own recent work, Giordano and Broderick [2023], being a notable exception).

Importantly, the local weights can be automatically computed with a small library built on top of existing open–source software which is commonly used for survey analysis [Lopez-Martin et al., 2022]. I have already implemented a similar package for a different style of sensitivity analysis [Broderick et al., 2020, Giordano, 2024], and we expect to be able to release open–source software relatively quickly.

References

AABI. 6th symposium on advances in approximate Bayesian inference, 2024. URL `https://approximateinference.org/`.

Eli B., Avi F., and Erin H. Multilevel calibration weighting for survey data, 2021.

D. Blei, A. Kucukelbir, and J. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.

T. Broderick, R. Giordano*, and R. Meager*. An automatic finite-sample robustness metric: Can dropping a little data change conclusions? *arXiv preprint arXiv:2011.14999*, 2020. *Following conventions in econometrics, the authors are listed alphabetically. Giordano and Meager are equal contribution primary authors.

K. Cranmer, J. Brehmer, and G. Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020.

A. Gelman. Poststratification into many categories using hierarchical logistic regression. *Survey Methodology*, 23:127, 1997.

A. Gelman. Struggles with survey weighting and regression modeling. 2007.

A. Gelman and M. Heidemanns. The Economist: Forecasting the US elections., 2020. URL `https://projects.economist.com/us-2020-forecast/president`. Data and model accessed Oct., 2020.

R. Giordano. Z-estimator approximate maximum influence: zaminfluence, 2024. URL `https://github.com/rgiordan/zaminfluence`.

R. Giordano and T. Broderick. The Bayesian infinitesimal jackknife for variance. *arXiv preprint arXiv:2305.06466*, 2023.

R. Giordano, W. Stephenson, R. Liu, M. I. Jordan, and T. Broderick. A Swiss army infinitesimal jackknife. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1139–1147. PMLR, 2019.

F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel. *Robust statistics: The approach based on influence functions*. Wiley-Interscience; New York, 1986.

F. Li, K. Morgan, and A. Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.

J. Lopez-Martin, J. Phillips, and A. Gelman. Multilevel regression and poststratification case studies, 2022. URL `https://bookdown.org/jl5522/MRP-case-studies/`.

R. Meager. Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, 11(1):57–91, 2019.

R. Mises. On the asymptotic distribution of differentiable statistical functions. *The Annals of Mathematical Statistics*, 18(3):309–348, 1947.

G. Papamakarios, E. Nalisnick, D. Rezende, S. Mohamed, and B. Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57): 1–64, 2021.

J. Pritchard, M. Stephens, and P. Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959, 2000.

J. Regier, K. Fischer, K. Pamnany, A. Noack, J. Revels, M. Lam, S. Howard, R. Giordano, D. Schlegel, and J. McAuliffe. Cataloging the visible universe through Bayesian inference in Julia at petascale. *Journal of Parallel and Distributed Computing*, 127:89–104, 2019.

S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman. Validating Bayesian inference algorithms with simulation-based calibration. *arXiv preprint arXiv:1804.06788*, 2018.